

Connecting via Winsock to STN

Welcome to STN International! Enter x:x

LOGINID:ssspta1649jxm

PASSWORD:

TERMINAL (ENTER 1, 2, 3, OR ?):2

\* \* \* \* \* Welcome to STN International \* \* \* \* \*

NEWS	1		Web Page URLs for STN Seminar Schedule - N. America
NEWS	2		"Ask CAS" for self-help around the clock
NEWS	3	May 12	EXTEND option available in structure searching
NEWS	4	May 12	Polymer links for the POLYLINK command completed in REGISTRY
NEWS	5	May 27	New UPM (Update Code Maximum) field for more efficient patent SDIs in CAPlus
NEWS	6	May 27	CAPlus super roles and document types searchable in REGISTRY
NEWS	7	Jun 28	Additional enzyme-catalyzed reactions added to CASREACT
NEWS	8	Jun 28	ANTE, AQUALINE, BIOENG, CIVILENG, ENVIROENG, MECHENG, and WATER from CSA now available on STN(R)
NEWS	9	Jul 12	BEILSTEIN enhanced with new display and select options, resulting in a closer connection to BABS
NEWS	10	Jul 30	BEILSTEIN on STN workshop to be held August 24 in conjunction with the 228th ACS National Meeting
NEWS	11	AUG 02	IFIPAT/IFIUDB/IFICDB reloaded with new search and display fields
NEWS	12	AUG 02	CAPlus and CA patent records enhanced with European and Japan Patent Office Classifications
NEWS	13	AUG 02	STN User Update to be held August 22 in conjunction with the 228th ACS National Meeting
NEWS	14	AUG 02	The Analysis Edition of STN Express with Discover! (Version 7.01 for Windows) now available
NEWS	15	AUG 04	Pricing for the Save Answers for SciFinder Wizard within STN Express with Discover! will change September 1, 2004
NEWS EXPRESS		JULY 30	CURRENT WINDOWS VERSION IS V7.01, CURRENT MACINTOSH VERSION IS V6.0c(ENG) AND V6.0Jc(JP), AND CURRENT DISCOVER FILE IS DATED 11 AUGUST 2004
NEWS HOURS			STN Operating Hours Plus Help Desk Availability
NEWS INTER			General Internet Information
NEWS LOGIN			Welcome Banner and News Items
NEWS PHONE			Direct Dial and Telecommunication Network Access to STN
NEWS WWW			CAS World Wide Web Site (general information)

Enter NEWS followed by the item number or name to see news on that specific topic.

All use of STN is subject to the provisions of the STN Customer agreement. Please note that this agreement limits use to scientific research. Use for software development or design or implementation of commercial gateways or other similar uses is prohibited and may result in loss of user privileges and other penalties.

\* \* \* \* \* STN Columbus \* \* \* \* \*

FILE 'HOME' ENTERED AT 13:04:37 ON 17 AUG 2004

=> file medline biosis embase caplus

COST IN U.S. DOLLARS	SINCE FILE ENTRY	TOTAL SESSION
FULL ESTIMATED COST	0.21	0.21

FILE 'MEDLINE' ENTERED AT 13:04:53 ON 17 AUG 2004

FILE 'BIOSIS' ENTERED AT 13:04:53 ON 17 AUG 2004  
COPYRIGHT (C) 2004 BIOLOGICAL ABSTRACTS INC.(R)

FILE 'EMBASE' ENTERED AT 13:04:53 ON 17 AUG 2004  
COPYRIGHT (C) 2004 Elsevier Inc. All rights reserved.

FILE 'CAPLUS' ENTERED AT 13:04:53 ON 17 AUG 2004  
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.  
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.  
COPYRIGHT (C) 2004 AMERICAN CHEMICAL SOCIETY (ACS)

=> s luo ying/au  
L1 219 LUO YING/AU

=> s mancebo helena/au  
L2 6 MANCEBO HELENA/AU

=> s syk (s) ubp (s) protein  
L3 3 SYK (S) UBP (S) PROTEIN

=> dup rem l3  
PROCESSING COMPLETED FOR L3  
L4 3 DUP REM L3 (0 DUPLICATES REMOVED)

=> d l4 total ibib kwic

L4 ANSWER 1 OF 3 BIOSIS COPYRIGHT 2004 BIOLOGICAL ABSTRACTS INC. on STN  
ACCESSION NUMBER: 2004:237339 BIOSIS  
DOCUMENT NUMBER: PREV200400237391  
TITLE: **SYK-UBP proteins**,  
compositions and methods of use.  
AUTHOR(S): Luo, Ying [Inventor, Reprint Author]; Mancebo, Helena  
[Inventor]  
CORPORATE SOURCE: San Bruno, CA, USA  
ASSIGNEE: Rigel Pharmaceuticals, Inc.  
PATENT INFORMATION: US 6709839 March 23, 2004  
SOURCE: Official Gazette of the United States Patent and Trademark  
Office Patents, (Mar 23 2004) Vol. 1280, No. 4.  
<http://www.uspto.gov/web/menu/patdata.html>. e-file.  
ISSN: 0098-1133 (ISSN print).  
DOCUMENT TYPE: Patent  
LANGUAGE: English  
ENTRY DATE: Entered STN: 28 Apr 2004  
Last Updated on STN: 28 Apr 2004

TI **SYK-UBP proteins**, compositions and methods  
of use.

IT Major Concepts  
Methods and Techniques; Pharmacology

IT Diseases  
disease: disease-miscellaneous  
Disease (MeSH)

IT Chemicals & Biochemicals  
**SYK-UBP proteins**: diagnostic-drug;  
nucleic acids; polypeptides

L4 ANSWER 2 OF 3 CAPLUS COPYRIGHT 2004 ACS on STN  
ACCESSION NUMBER: 2003:376376 CAPLUS  
DOCUMENT NUMBER: 138:364551

TITLE: Novel human SYK kinase-associated cell cycle protein sequence and modulation of leukocyte activation, compositions and methods of use

INVENTOR(S): Wong, Brian; Fu, Chong Alan; Mancebo, Helena; Zhou, Xiulan X. Z.

PATENT ASSIGNEE(S): USA

SOURCE: U.S. Pat. Appl. Publ., 88 pp., Cont.-in-part of U.S. Ser. No. 39,761.  
CODEN: USXXCO

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 3

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 2003092605	A1	20030515	US 2002-112674	20020329
US 2003036107	A1	20030220	US 2001-39761	20011019
PRIORITY APPLN. INFO.:			US 2001-280698P	P 20010329
			US 2001-39761	A2 20011019
			US 1999-404967	A 19990924
			WO 2000-US26338	W 20000925

AB The present invention provides compns. and methods for modulating leukocyte activation. The invention provides **protein** and cDNA sequences for novel human **SYK** kinase-associated cell cycle **proteins SYK-UBP**. Nucleic acids encoding proteins and proteins so encoded which are capable of modulating leukocyte activation are provided. Compns. and methods for the treatment of disorders related to leukocyte dysfunction or dysregulation are also provided. Prophylactics and methods for the prevention of such disorders are also provided. Also provided are compns. and methods for diagnostic and prognostic determination of such disorders. Further provided are assays for the identification of bioactive agents capable of modulating leukocyte activation.

IT 331009-92-8D, **Protein SYK-UBP** (human isoform 1), Subfragments are claimed

RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)  
(amino acid sequence; novel human **SYK** kinase-associated cell cycle **protein** sequence and modulation of leukocyte activation, compns. and methods of use)

L4 ANSWER 3 OF 3 CAPLUS COPYRIGHT 2004 ACS on STN

ACCESSION NUMBER: 2001:228922 CAPLUS

DOCUMENT NUMBER: 134:247997

TITLE: Novel **SYK** kinase-associated cell cycle **proteins SYK-UBPs**, compositions and methods of use

INVENTOR(S): Luo, Ying; Mancebo, Helena

PATENT ASSIGNEE(S): Rigel Pharmaceuticals, Inc., USA

SOURCE: PCT Int. Appl., 63 pp.  
CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 3

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2001021654	A2	20010329	WO 2000-US26338	20000925
WO 2001021654	A3	20020307		
W: AU, CA, JP, US				
RW: AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL,				

	PT, SE				
US 6709839	B1	20040323	US 1999-404967		19990924
AU 2000078328	A5	20010424	AU 2000-78328		20000925
EP 1218505	A2	20020703	EP 2000-968406		20000925
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI, CY					
JP 2003512820	T2	20030408	JP 2001-525227		20000925
US 2003036107	A1	20030220	US 2001-39761		20011019
PRIORITY APPLN. INFO.:			US 1999-404967	A2	19990924
			WO 2000-US26338	W	20000925
			US 2001-280698P	P	20010329

TI Novel **SYK** kinase-associated cell cycle **proteins**  
**SYK-UBPs**, compositions and methods of use

AB The invention provides **protein** and cDNA sequences for novel human **SYK** kinase-associated cell cycle **proteins** **SYK-UBPs**, which have an effect on or are related to the cell cycle. Also provided herein are vectors and host cells comprising those nucleic acid sequences, chimeric polypeptide mols. comprising the polypeptides of the present invention fused to heterologous polypeptide sequences, antibodies which bind to the polypeptides of the present invention and to methods for producing the polypeptides of the present invention. Further provided by the present invention are methods for identifying novel compns. which mediate cell cycle bioactivity, and the use of such compns. in diagnosis and treatment of disease.

IT **Proteins**, specific or class  
 RL: BPN (Biosynthetic preparation); BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); PROC (Process); USES (Uses)  
 (SYK-UBP, of human; novel **SYK** kinase-associated cell cycle **proteins** **SYK-UBPs**, compns. and methods of use)

IT cDNA sequences  
 (for **protein** **SYK-UBP**, of human; novel **SYK** kinase-associated cell cycle **proteins** **SYK-UBPs**, compns. and methods of use)

IT Gene, animal  
 RL: BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); USES (Uses)  
 (for **protein** **SYK-UBP**, of human; novel **SYK** kinase-associated cell cycle **proteins** **SYK-UBPs**, compns. and methods of use)

IT Diagnosis  
 (mol.; novel **SYK** kinase-associated cell cycle **proteins** **SYK-UBPs**, compns. and methods of use)

IT Antibodies  
 RL: THU (Therapeutic use); BIOL (Biological study); USES (Uses)  
 (monoclonal, to **protein** **SYK-UBP**; novel **SYK** kinase-associated cell cycle **proteins** **SYK-UBPs**, compns. and methods of use)

IT Cell cycle  
 Drug screening  
 Gene therapy  
 Molecular cloning  
 Nucleic acid hybridization  
 (novel **SYK** kinase-associated cell cycle **proteins** **SYK-UBPs**, compns. and methods of use)

IT Primers (nucleic acid)  
 Probes (nucleic acid)  
 RL: ARG (Analytical reagent use); ANST (Analytical study); USES (Uses)  
 (novel **SYK** kinase-associated cell cycle **proteins** **SYK-UBPs**, compns. and methods of use)

IT **Protein** sequences  
 (of **protein** **SYK-UBP**, of human; novel **SYK** kinase-associated cell cycle **proteins** **SYK-**

UBPs, compns. and methods of use)

IT Antibodies  
 RL: THU (Therapeutic use); BIOL (Biological study); USES (Uses)  
 (to **protein SYK-UBP**; novel **SYK**  
 kinase-associated cell cycle **proteins SYK-UBPs**  
 , compns. and methods of use)

IT **Protein motifs**  
 (ubiquitin-associated binding domain in **protein SYK-UBP**; novel **SYK** kinase-associated cell cycle **proteins SYK-UBPs**, compns. and methods of use)

IT **Protein motifs**  
 (ubiquitin-specific protease domain in **protein SYK-UBP**; novel **SYK** kinase-associated cell cycle **proteins SYK-UBPs**, compns. and methods of use)

IT 331009-92-8P, **Protein SYK-UBP** (human isoform 1) 331009-93-9P, **Protein SYK-UBP** (human isoform 2)  
 RL: BPN (Biosynthetic preparation); BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); PROC (Process); USES (Uses)  
 (amino acid sequence; novel **SYK** kinase-associated cell cycle **proteins SYK-UBPs**, compns. and methods of use)

IT 60267-61-0, Ubiquitin 138674-26-7, **SYK** kinase  
 RL: BPR (Biological process); BSU (Biological study, unclassified); BIOL (Biological study); PROC (Process)  
 (cell cycle **protein SYK-UBP** interacting with; novel **SYK** kinase-associated cell cycle **proteins SYK-UBPs**, compns. and methods of use)

IT 331009-94-0 331009-95-1  
 RL: BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); USES (Uses)  
 (nucleotide sequence; novel **SYK** kinase-associated cell cycle **proteins SYK-UBPs**, compns. and methods of use)

IT 331011-23-5  
 RL: PRP (Properties)  
 (unclaimed **protein** sequence; novel **SYK** kinase-associated cell cycle **proteins SYK-UBPs**, compns. and methods of use)

L Number	Hits	Search Text	DB	Time stamp
1	27	luo-ying.in.	USPAT; US-PGPUB; EPO; JPO; DERWENT	2004/08/17 11:55
2	4	mancebo-helena.in.	USPAT; US-PGPUB; EPO; JPO; DERWENT	2004/08/17 11:56
3	3	syk-ubp same protein	USPAT; US-PGPUB; EPO; JPO; DERWENT	2004/08/17 11:56

## 10088960 Results

SEQ ID NO: 2

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	5473	100.0	1055	22	AAB74491	Human SYK kinase b
2	5460	99.7	1055	22	AAB31546	A human ubiquitin
3	5460	99.7	1055	22	AAB31552	A human ubiquitin
4	5453	99.6	1075	22	AAB74492	Human SYK kinase b
5	5430	99.2	1087	22	AAB31550	A human ubiquitin
6	5430	99.2	1087	22	AAB31556	A human ubiquitin
7	4308	78.7	834	22	AAB74493	Human SYK kinase b
8	3741	68.3	720	20	AAAY45343	Human secreted pro
9	344	6.3	71	21	AAG03009	Human secreted pro
10	318.5	5.8	1264	14	AAR36730	Ubiquitin-specific
11	224	4.1	2547	19	AAW81511	Drosophila fat fac
12	215	3.9	912	22	AAB75607	Human cancer assoc
13	214.5	3.9	2954	20	AAAY01632	Amino acid sequenc
14	213.5	3.9	1517	21	AAG38482	Arabidopsis thalia
15	213.5	3.9	1543	21	AAG38481	Arabidopsis thalia

## RESULT 8

AAAY45343

ID AAY45343 standard; Protein; 720 AA.

XX

AC AAY45343;

XX

DT 23-NOV-1999 (first entry)

XX

DE Human secreted protein fragment encoded from gene 23.

XX

KW Secreted protein; human; gene therapy; diagnosis; treatment; cancer;  
 KW protein therapy; tumor; neurodegenerative disorder; blood disorder; AIDS;  
 KW developmental abnormality; leukemia; immune system; autoimmune disease;  
 KW hepatic disease; renal disease; inflammation; allergy; schizophrenia;  
 KW Alzheimer's disease; cognitive disorder; arthritis; infection; psoriasis;  
 KW transplant rejection; diabetes; asthma; sepsis; acne; metabolic disorder;  
 KW cardiovascular disorder; food additive; preservative.

XX

OS Homo sapiens.

XX

PN W09946289-A1.

XX

PD 16-SEP-1999.

XX

PF 11-MAR-1999; 99WO-US05721.

XX

PR 12-MAR-1998; 98US-0077686.

PR 12-MAR-1998; 98US-0077687.

PR 12-MAR-1998; 98US-0077696.

PR 12-MAR-1998; 98US-0077714.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

XX

PI Ruben SM, Ferrie AM, Rosen CA, Florence C, Young PE, Yu G, Ni J;

XX

DR WPI; 1999-551363/46.

DR N-PSDB; AAZ27255.

XX

PT New isolated human genes, useful for diagnosis and treatment of, e.g.  
 PT cancers -

XX

PS Disclosure; 214-216; 306pp; English.

XX

CC This invention describes novel human genes and the secreted proteins  
 CC they encode. The polynucleotides and their corresponding secreted  
 CC polypeptides are useful for preventing, treating or ameliorating medical

CC conditions e.g. by protein or gene therapy. Also pathological conditions  
 CC can be diagnosed by determining the amount of the new polypeptides in a  
 CC sample or by determining the presence of mutations in the new  
 CC polynucleotides. Specific uses are described for each of the  
 CC polynucleotides of the invention, based on which tissues they are most  
 CC highly expressed in, and include developing products for the diagnosis or  
 CC treatment of cancer, tumors, neurodegenerative disorders, developmental  
 CC abnormalities, blood disorders, leukemias, diseases of the immune system,  
 CC autoimmune diseases, hepatic and renal disease, inflammation, allergies,  
 CC Alzheimer's and cognitive disorders, schizophrenia, arthritis,  
 CC infections, AIDS, transplant rejection, diabetes, asthma, sepsis, acne,  
 CC psoriasis, cardiovascular disorders, and metabolic disorders. The  
 CC polypeptides or polynucleotides can also be used as food additives or  
 CC preservatives. The polypeptides are also useful for identifying their  
 CC binding partners. AAY45290-Y45389 represent human secreted protein  
 CC fragments described in the method of the invention.

XX

SQ Sequence 720 AA;

Query Match 68.3%; Score 3741; DB 20; Length 720;  
 Best Local Similarity 99.7%; Pred. No. 2.5e-277;  
 Matches 718; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 336 CLEAAMIEGIEIESLHSENSGKSGQEHWFTGLPPVLTFXLSRFEFNQALGRPEKIHNKLEF 395  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 1 cleaamiegieieslhsensgksgqehwftelppvltfelsrfefnqalgrpekihnklef 60

Qy 396 PQVLYLDRYMHRNREITRIKREEIKRLKDYLTVLQQRRLERYLSYSGGPKRFPPLVDVLQYA 455  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 61 pqvlyldrymhrnreitrikreeikrlkdytlvlqqrllerylsyggpkrfplvdvlqya 120

Qy 456 LEFASSKPVCTSPVDDIDASSPPSGSIPSQTLPTSTTEQQGALSSSELPSTSPSSVAAISSR 515  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 121 lefasskpvtspvddidassppsgsipsqtlpstteqqgalsselpstspssvaaissr 180

Qy 516 SVIHKPFTQSRIPDLPMPAPRHITTEELSULESLHWRTEIENDTRDLQESISRIHR 575  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 181 svihkpftqsrippdlpmhpaprHITEELSULESLHWRTEIENDTRDLQESISRIHR 240

Qy 576 TIELMYSKSMIQVYPYRLHAVLVHEGGANAGHYWAYIFDHRESRWKYNIDIAVTKSSWEE 635  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 241 tielmysksmiqvpyrlhavlvehgganaghywayifdhresrwmkyndiavtksswee 300

Qy 636 LVRDSFGGYRNASAYCLMYINDKAQFLIQEEFNKETGQPLVGIETLPPDLRDFVEEDNQR 695  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 301 lvrdsfgyrnasayclmyindkaqfliqeefnketgqplvgietlppdlrdfveednqr 360

Qy 696 FEKELEEWDAQLAQKALQEKLLASQKLRESETSVTTAQAAGDPEYLEQPSRSDFSKHLKE 755  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 361 fekeleewdaqlaqkalqekllasqklresetsvttaqaagdpeyleqpsrsdfskhlke 420

Qy 756 ETIQIITKASHEHEDKSPETVLQSAIKLEYARLVKLAQEDTPPETDYRLHHVVVYFIQNG 815  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 421 etiqiitkashehedkspetvlqsaikleyarlvklaqedtppetdyrlhhvvvyfiqng 480

Qy 816 APKKIIEKTLLEQFGDRNLSFDERCHNIMKVAQAKLEMIPPEEVNLEEYEEWHQDYRKFR 875  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 481 apkkiiektlleqfgdrnlsfderchnimkvaqaklemippeevnleeyeewhqdyrkfr 540

Qy 876 ETTMYLIIGLENFQRESYIDSLFLICAYQNNKELLSKGLYRGHDEELISHYRRECLLKL 935  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 541 ettmyliiglenfqresyidsllflicayqnnkellskglyrghdeelishyrrecllkl 600

Qy 936 NEQAAELFESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEEKDILAVEDMNRNRWCSYLQ 995  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 601 neqaaelfesgedrevnnngliimnefivpflplllvdemeekdilavedmnrnrwcsylq 660

Qy 996 EMEPHLQEKLTDFLPKLLDCSMEIKSFHEPPKLPSTHELCECFARIMLSLSRTPADGR 1055  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 661 emephqlqekltdfllpklldcsmeksfheppklpsysthelcerfarimlslsrtpadgr 720



Result No.	Score	Query Match	Length	DB	ID	Description
1	327.5	6.0	1264	1	US-07-789-915A-6	Sequence 6, Appli
2	327.5	6.0	1264	1	US-08-005-002C-6	Sequence 6, Appli
3	327.5	6.0	1264	1	US-08-487-203A-6	Sequence 6, Appli
4	224	4.1	2547	3	US-09-058-489-35	Sequence 35, Appl
5	210	3.8	2555	3	US-09-058-489-36	Sequence 36, Appl
6	186.5	3.4	1786	4	US-08-973-462-8	Sequence 8, Appli
7	178.5	3.3	1164	4	US-08-923-992A-10	Sequence 10, Appl
8	175	3.2	1312	2	US-08-687-080-51	Sequence 51, Appl
9	174	3.2	1312	2	US-08-592-126-148	Sequence 148, App
10	169	3.1	809	1	US-07-789-915A-4	Sequence 4, Appli
11	169	3.1	809	1	US-08-005-002C-4	Sequence 4, Appli

Result No.	Score	Query Match	Length	DB	ID	Description
1	493.5	9.0	1302	2	T23236	hypothetical prote
2	323.5	5.9	1272	2	S60999	ubiquitin-specific
3	313.5	5.7	1108	2	T41188	probable ubiquitin
4	310	5.7	1221	2	T30529	ubiquitin carboxyl
5	287	5.2	1130	2	T23104	hypothetical prote
6	287	5.2	1133	2	T23103	hypothetical prote
7	268.5	4.9	3187	2	JC5837	364K Golgi complex
8	263.5	4.8	1230	2	T19899	hypothetical prote
9	261.5	4.8	1230	2	S53974	hypothetical prote
10	239.5	4.4	1430	2	T21910	hypothetical prote
11	232	4.2	1292	2	T24559	hypothetical prote

Result No.	Score	Query Match	Length	DB	ID	Description
1	3836	70.1	737	4	Q9H9W1	Q9h9w1 homo sapien
2	2741	50.1	523	4	Q9HA22	Q9ha22 homo sapien
3	1895	34.6	757	4	Q9P213	Q9p213 homo sapien
4	301	5.5	1124	10	Q9SF08	Q9sf08 arabidopsis
5	296	5.4	1248	5	Q9VRP1	Q9vrp1 drosophila
6	294	5.4	1126	10	Q9FG10	Q9fg10 arabidopsis
7	293.5	5.4	1116	10	Q9FPT1	Q9fpt1 arabidopsis
8	288.5	5.3	599	10	Q9FU99	Q9fu99 arabidopsis
9	287	5.2	1130	5	O45624	O45624 caenorhabdi
10	287	5.2	1133	5	O45623	O45623 caenorhabdi
11	283.5	5.2	1141	3	Q9P3U0	Q9p3u0 schizosacch
12	268.5	4.9	3187	11	Q63714	Q63714 rattus norv
13	264.5	4.8	931	4	Q9HCH8	Q9hch8 homo sapien
14	264.5	4.8	1129	3	Q9UTT1	Q9utt1 schizosacch
15	263.5	4.8	1230	5	Q93371	Q93371 caenorhabdi

SEQ ID NO: 4

Result No.	Score	Query Match	Length	DB	ID	Description
1	5579	100.0	1075	22	AAB74492	Human SYK kinase b
2	5453	97.7	1055	22	AAB74491	Human SYK kinase b
3	5440	97.5	1055	22	AAB31546	A human ubiquitin
4	5440	97.5	1055	22	AAB31552	A human ubiquitin
5	5410	96.9	1087	22	AAB31550	A human ubiquitin
6	5410	96.9	1087	22	AAB31556	A human ubiquitin
7	4308	77.2	834	22	AAB74493	Human SYK kinase b
8	3721	66.7	720	20	AA45343	Human secreted pro
9	344	6.2	71	21	AAG03009	Human secreted pro
10	318.5	5.7	1264	14	AAR36730	Ubiquitin-specific
11	224	4.0	2547	19	AAW81511	Drosophila fat fac

RESULT 8  
 AAY45343  
 ID AAY45343 standard; Protein; 720 AA.  
 XX  
 AC AAY45343;  
 XX  
 DT 23-NOV-1999 (first entry)  
 XX  
 DE Human secreted protein fragment encoded from gene 23.  
 XX  
 KW Secreted protein; human; gene therapy; diagnosis; treatment; cancer;  
 KW protein therapy; tumor; neurodegenerative disorder; blood disorder; AIDS;  
 KW developmental abnormality; leukemia; immune system; autoimmune disease;  
 KW hepatic disease; renal disease; inflammation; allergy; schizophrenia;  
 KW Alzheimer's disease; cognitive disorder; arthritis; infection; psoriasis;  
 KW transplant rejection; diabetes; asthma; sepsis; acne; metabolic disorder;  
 KW cardiovascular disorder; food additive; preservative.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9946289-A1.  
 XX  
 PD 16-SEP-1999.  
 XX  
 PF 11-MAR-1999; 99WO-US05721.  
 XX  
 PR 12-MAR-1998; 98US-0077686.  
 PR 12-MAR-1998; 98US-0077687.  
 PR 12-MAR-1998; 98US-0077696.  
 PR 12-MAR-1998; 98US-0077714.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Ruben SM, Ferrie AM, Rosen CA, Florence C, Young PE, Yu G, Ni J;  
 XX  
 DR WPI; 1999-551363/46.  
 DR N-PSDB; AAZ27255.  
 XX  
 PT New isolated human genes, useful for diagnosis and treatment of, e.g.  
 PT cancers -  
 XX  
 PS Disclosure; 214-216; 306pp; English.  
 XX  
 CC This invention describes novel human genes and the secreted proteins  
 CC they encode. The polynucleotides and their corresponding secreted  
 CC polypeptides are useful for preventing, treating or ameliorating medical  
 CC conditions e.g. by protein or gene therapy. Also pathological conditions  
 CC can be diagnosed by determining the amount of the new polypeptides in a  
 CC sample or by determining the presence of mutations in the new  
 CC polynucleotides. Specific uses are described for each of the  
 CC polynucleotides of the invention, based on which tissues they are most  
 CC highly expressed in, and include developing products for the diagnosis or  
 CC treatment of cancer, tumors, neurodegenerative disorders, developmental  
 CC abnormalities, blood disorders, leukemias, diseases of the immune system,  
 CC autoimmune diseases, hepatic and renal disease, inflammation, allergies,  
 CC Alzheimer's and cognitive disorders, schizophrenia, arthritis,  
 CC infections, AIDS, transplant rejection, diabetes, asthma, sepsis, acne,  
 CC psoriasis, cardiovascular disorders, and metabolic disorders. The  
 CC polypeptides or polynucleotides can also be used as food additives or  
 CC preservatives. The polypeptides are also useful for identifying their  
 CC binding partners. AAY45290-Y45389 represent human secreted protein  
 CC fragments described in the method of the invention.  
 XX  
 SQ Sequence 720 AA;

Query Match 66.7%; Score 3721; DB 20; Length 720;  
 Best Local Similarity 97.0%; Pred. No. 1.9e-278;  
 Matches 718; Conservative 0; Mismatches 2; Indels 20; Gaps 1;

Qy 336 CLEAAMIEGEIESLHSENSGKSGQEHWF TGLPPVLT FXLSRFEFNQALGRPEKIH NKLEF 395  
 |||  
 Db 1 cleaamiegeieslhsensgksgqehwftelppvltfelsrfe fnqalgrpekihnklef 60

Qy 396 PQVLYLD RYMHNRNREITRIKREEIKRLKDYLT V LQQRLERYLSYSGGPKRFPLVDVLQYA 455  
 |||  
 Db 61 pqvlyldrymhrnreitrikreeikrlkdy ltv lqqrlerylsyggpkrfplvdvlqya 120

Qy 456 LEFASSKPVCTSPVDDIDASSPPSGSIPSQTL PSTTEQQGALSSSELPSTSPSSVAAISSR 515  
 |||  
 Db 121 lefasskpvctspvddidassppsgsipsq tlpstteqqgalsselpstspssvaaissr 180

Qy 516 SVIHKPFTQSRIPDLPMPAPRHIT EEEELSVLESCLHRWRTEIENDTRDLQESISRIHR 575  
 |||  
 Db 181 svihkpf tqsrippdlpmhpaprhiteeelsvlesclhrwrteiendtrdlqesisrihr 240

Qy 576 TIELMYSKSMIQVPYRLHAVLVHEGQANAGHYWAYIFDHRESRWMKYNDIAVTKSSWEE 635  
 |||  
 Db 241 tielmysdksmiqvpyrlhavlvhegganaghywayifdhresrwmkyndiavtksswee 300

Qy 636 LVRDSFGGYRNASAYCLMYINDKAQFLIQEEFNKETGQPLVG IETLPPDLRDFVEEDNQR 695  
 |||  
 Db 301 lvrdsfggyrnasayclmyindkaqfliqee fnketgqplvgietlppdlrdfveednqr 360

Qy 696 FEKELEEWDAQLAQKALQEKL LSAQKLRESETSVTTAQAAGDPEYLEQPSRSDFSKHLKE 755  
 |||  
 Db 361 fekeleewdaq laqkalqekllsaqklre setsvttaqaagdpeyleqpsrsdfskhlke 420

Qy 756 ETIQIITKASHEHEDKSPETVLQSAIKLEYARLVKLAQEDTPPETDYRLHHVVVYFIQNO 815  
 |||  
 Db 421 etiqiitkashehedkspetvlqsaikleyarlvklaqedtppetdyrlhhvvvyfiqnq 480

Qy 816 APKKIIEKTLLEQFGDRNLSFDERCHNIMKVAQAKLEMIKPEEVNLEEYEEWHQDYRKFR 875  
 |||  
 Db 481 apkkiektlleqfgdrnlsfderchnimkvaqaklemikpeevnleeyee whqdyrkfr 540

Qy 876 ETTMYLIIGLENFQRESYIDSLFLICAYQNNKELLSKGLYRGHDEELISHYRRECLLIL 935  
 |||  
 Db 541 ettmyliiglenfqresyidsllflicayqnnkellskglyrghdeelishyrrecll-- 598

Qy 936 NLKRKQKPILFFFLHCICKLNEQA AELFESGEDREVNNGLIIMNEFIVPFLPLLVD EME 995  
 |||  
 Db 599 -----klneqaaelfesgedrevnn gliimnefivpflplllvdeme 640

Qy 996 EKDILAVEDMRNRWCSYLGQEMEPHLQEKL TDFLPKLLDCSMEIKSFHEPPKLPSYS THE 1055  
 |||  
 Db 641 ekdilavedmnrwcsylgqemephlqekl tdf lpklldcsmeksfheppklpsys the 700

Qy 1056 LCERFARIMLSLSRTPADGR 1075  
 |||  
 Db 701 lcerfarimlslsrtpadgr 720

RESULT 10

AAR36730

ID AAR36730 standard; Protein; 1264 AA.

XX

AC AAR36730;

XX

DT 26-AUG-1993 (first entry)

XX

DE Ubiquitin-specific protease, UBP2.

XX

KW Fusion protein; ubiquitin-methionine-beta-galactosidase; UBP2; UBP3;

KW expression vector; ubiquitin-specific; protease; UBP1; in vitro;

KW transformation; Ub-met-beta-Gal; in vivo.

XX

OS Synthetic.

XX

PN WO9309235-A.

XX  
 PD 13-MAY-1993.  
 XX  
 PF 06-NOV-1992; 92WO-US09562.  
 XX  
 PR 08-NOV-1991; 91US-0789915.  
 XX  
 PA (MASI ) MASSACHUSETTS INST TECHNOLOGY.  
 XX  
 PI Baker RT, Tobias JW, Varshavsky A;  
 XX  
 DR WPI; 1993-167700/20.  
 DR N-PSDB; AAQ41289.  
 XX  
 PT Ubiquitin-specific protease(s) - specifically cleave fusion  
 PT proteins regardless of size, at C-terminus of ubiquitin portion  
 XX  
 PS Claim 6; Page 47-54; 78pp; English.  
 XX  
 CC The sequences given in AAR36729-31 represent the ubiquitin-specific  
 CC proteases UBP1, UBP2, and UBP3. These proteases specifically cleave  
 CC the 120 kD fusion protein ubiquitin-methionine-beta-galactosidase  
 CC (Ub-met-beta-Gal) (see also AAR36728). The proteases UBP1 and UBP2  
 CC demonstrate activity both in vivo and in vitro, whereas UBP3 is only  
 CC active in vivo. These proteases have been shown to specifically  
 CC cleave Ub-met-beta-Gal at the C-terminus of the ubiquitin moiety.  
 CC The proteases can be used to deubiquinate fusion proteins in vivo,  
 CC therefore prokaryotic cells having an expression vector one of these  
 CC proteases can be further transformed with an expression vector  
 CC encoding a ubiquitin fusion protein. Such cells will then produce  
 CC a deubiquinated product having a predetermined N-terminal amino acid  
 CC residue.  
 XX  
 SQ Sequence 1264 AA;

Query Match 5.7%; Score 318.5; DB 14; Length 1264;  
 Best Local Similarity 19.6%; Pred. No. 1.8e-15;  
 Matches 170; Conservative 121; Mismatches 240; Indels 337; Gaps 33;

Qy 4 EQNVLQQSAAQKHQQTTF-----LNQLREITGIN---DTQILQQALKDSNGNLELAV 51  
 :|::| | : | :| | | : : | : | : | : | :  
 Db 502 kqdiigqealenallmfkinpkecniselneatlslsykyetsnksqvtshltnlknal 561  
  
 Qy 52 AFLTAKNAKTPQQEETTYQTALPGNDRYISVGSQADTNVIDLTGDD----- 98  
 | | : : : | | : : | | | : | :  
 Db 562 rll---akytsdklkyfyvdhep---yralsqaydtlsidesvdediiktaysvkind 613  
  
 Qy 99 ----KDDLQR---TIALS-----LAESNRAFRET----- 120  
 | | | | | : | : | : | : | : | :  
 Db 614 spglklcdralytlaiskrslldfnflteecpqfsnyygppekllqvnenasdetilkif 673  
  
 Qy 121 -----GITDEEQAI---SRVLEASIAENKACL-----K RTPTEVWRDSRNPY 159  
 : : : | : : : | | : : | : | : | : | :  
 Db 674 kqkwfdenvyepdqflilraaltkislernstltnflttgtidpnsllpenw----- 726  
  
 Qy 160 DRKRQDKAPVGLKNVGNCTCWFSAVIQSLFNLLFRRRLVLNY-KPPSNAQD----- 208  
 | | : | | | : : : | : : | | | | | | | |  
 Db 727 -----ptginnigntcylnsllyqyfsiaplrryvleyqktvenfndhlsnsghirr 778  
  
 Qy 209 -----LPRNQKEHRNLPFMRELRYLFALLVGTKRKYVDPSRAV----- 246  
 : | : | : : | : : | | : | : | : | : :  
 Db 779 iggrqisrgeve-rsiqfiyqlrnlfyamvhtercvtpskelaylafapsnvevefeve 837  
  
 Qy 247 -----EILKDAF-----KSNDSSQQQD 262  
 | | | | | : | : | : | : | : | :  
 Db 838 gnkvvdtgtvlsdskketddafsttkidtslidlemedglnsdvgtgdanrkknendae 897  
  
 Qy 263 VSE-----FTHKLLDWLEDAFQMKAEET-----DEEK 290  
 | | | | | : | : | : | : | : | :  
 Db 898 vsenedttgtlsptrvakissdqlenalemgrqqdvtecignvlfqiesgsepirydedn 957

Qy	291	PKNPMV-ELFYG-----RFLAVGVLEGKKFENTEMFGQYPLQVNG	329
		: :  :       : :   :	
Db	958	eqydlvkqlfygttkgsivplsatnkvrtkverflslli-----nigdhp----	1002
Qy	330	FKDLHECLEAAMIEGEIESLHSENSGKSGQEHWFTGLPPVLTFXLSRFEFNQALGRPEKI	389
		::: : :       :   :  :  :  :  :	
Db	1003	-kdiyda fdsyfk d---eyltmeeygdvirtvavttftptilqvqigrvydrerlmpfks	1058
Qy	390	HNKLEFPQVLYLDRYMHRNREITRIKREEIKRLKDYLTVLQQLRLERYLSY-GSGPKRFPL	448
		:  :  :  : : :  : :  :  :  :  :	
Db	1059	ieplpfkeviymdryadtenplllakketeemqgklkvmmknrqrellsrddsgltr---	1115
Qy	449	VDVLQYALEFASSKPVCTSPVDDIDASSPPSGSIPSQTLPTSTTEQQGALSSELPSTSPSS	508
		:  :  :  :  :  :  :  :  :  :	
Db	1116	kda flesikllesdtikktpl-kiea-----andviktlrnn	1151
Qy	509	VAAISSRSVIHKPFTQSRIPDPLMPHPAPRHITEELS SVLESCLHRWTEIENDTRDLQE	568
		:  :  :  :  :  :  :  :  :	
Db	1152	vqni-----dnel-----mklyndinslee	1171
Qy	569	SISRIRHTIELMYSDKSMIQVPYRLHAVLVHEGQANAGHYWAYIFD-HRESRWKMYNDIA	627
		:  :  :  :  :  :  :  :  :  :	
Db	1172	kish-----qfdd--fkeygyslfsvfihrgeasyrhywiykdrnrngiwrkyndet	1222
Qy	628	VTKSSWEELVRDSFGGYRNASAYCLMYI	655
		:::   : :  :  :  :	
Db	1223	isevqeeevf--nfneqnatatpyflvyv	1248

A;Molecule type: DNA  
A;Residues: 1-1302 <WIL>  
A;Cross-references: EMBL:Z47811; PIDN:CAA87786.1; GSPDB:GN00020; CESP:K02C4.3  
A;Experimental source: clone K02C4  
C;Genetics:  
A;Gene: CESP:K02C4.3  
A;Map position: 2  
A;Introns: 10/1; 31/3; 84/3; 279/3; 464/1; 745/1; 957/2; 978/3; 1090/2; 1201/3; 1270/3

Query Match 8.7%; Score 485.5; DB 2; Length 1302;  
Best Local Similarity 23.2%; Pred. No. 1.2e-18;  
Matches 237; Conservative 139; Mismatches 353; Indels 293; Gaps 45;

```

Qy   90 NVIDLTGDDKDDLQRTIALSLAESNRAFRETGITDEEQAISRVLEASIAENKACLKRTPT 149
      | : :: : : : | | | | | | | | | | | | | | | | | | | | | |
Db  150 NISEISKKEEVEMQSAIQQLASS-----ASQNISR-----PT 182

Qy   150 EVWRDSRNPYDRKRQDKAPVGLKNVGNCTWFSAVIQSLFNLLFFRRLVNLKPPSNAQDL 209
      : | | | | | | | | | | | | | | | | | | | | | | | | :
Db  183 ML---MSNLEDMVRNPNFSTGLYNSGNTCWLNCLSQVLYSIPKFRSILYHCAPLSWHEQP 239

Qy   210 PRNQK----EHRNLPFMRELRYLFALLVGTKRKYVDPSRAVEILKDAFKSND-----SQQ 260
      | | : | : | | | | | | | | | | | | | | | | | | | |
Db  240 ITNVIENQQAELML--FRGLFAELQFSEMKYIEVGPLINMVDKLSKSSKGPSTIGTQ 297

Qy   261 QDVSEFTHKLLEDWLEDAF-----QMKAE--EETDEEK----- 290
      | | : | : | | | | | | | | | | | | | | | | | | | |
Db  298 QDATEMLTLIFDWLQRAFDAALHAQLNPEFSNVSDENLVISDSTTTAPNSDIIGAPPGY 357

Qy   291 -----PK---NPM-----VELFYGRFLAVGV-----L 309
      | | | | | | | | | | | | | | | | | | | | | | |
Db  358 NAANLSLPSSSHVDPKSTLNPYVNEKEPSSTPTSLFGTRSKTIEVNESMDTEAATSSNL 417

Qy   310 EGKKFEN-----TEMFGQ----- 322
      | | | | | | | | | | | | | | | | | | | | | | |
Db  418 PGNSVENHPNPAPEVDDNKAFCDKLKESFNNIFSSVCYTESVAEDGTVSVKSNVRNCP 477

Qy   323 --YPLQVNGFKDLHECLEAAMIEGEIESLHSENSGKSGQEHWFTGLPPVLTFXLSRFEFN 380
      : | | | | : | | | | | : | | : | | | | | | | |
Db  478 QFFQLQVT-YGNLHDALEAATFD-----HGLGNTASHVRNLYDPLPAVIFGLSRFSFN 530

Qy   381 QALGRPEKIHKNLEFPQVLYLDRYMHRNREITRIKREEIKRLKDYTLVLQQRRLERYLSYG 440
      : | : | | | | | | | | | | | | | | | | | | | |
Db  531 SNI--ESKLHDKFTFPKIIIFMDRYLKCNKEQLVQLRSHRELCRDSLSEVRAKLSGLRRYP 588

Qy   441 SGPKRFPLVDVLQYALEFASS-KPVCTSPVDDIDASSPPSGSIPSQTLPS---TTEQQGA 496
      | | | | | | | | | | : | : | : | | | | | | |
Db  589 QGNGEVRLEDSFQTVWQAVSNFREFVTFYL-----KVSQKTFFSREDAHENTAF 637

Qy   497 LSSELPTSPSSVAAISSRSV-----IHKPFTQSRIPDLPMPHAPRHITEEELSVLESC 551
      : | | | | | | | | | | : | | | | | : | | | |
Db  638 VGPLTPSTYQSSSDNCSSKFVKDGGKLFPTFTEGFFPGKAAF-----IETLQNM 686

Qy   552 LHRWRTEIENDTRLQESISRIHRTIELMYSKSMIQVPYRLHAVLVHEGQANAGHYWAY 611
      | : | | | | | | | | | | : | | | | | | | | | |
Db  687 LEALKTE----ERDCLAEEARLQEVIDQTYEVPQLQHKYELHAIIVHSGEANRGHYWTY 742

Qy   612 ----IFDHRESRWKYNIDIAVTKSSWEELVRDSFG-GYRNA-SAYCLMYINDKAQFLIQE 665
      | | | | | | | | | | : | : | | | | | | | | |
Db  743 KLKKSIDGLE-EWEKLNQADRVDWPKVESDSFGTGSRDAPSAYMLMYVRSDAEWLVSA 801

Qy   666 EFNKETGQPLVGIETLPPDLRDFVEEDNQRFEKELEEWDAQLAQKALQEKLLASQKLRES 725
      : | | | | | | | | | | : | : | | | | | | | |
Db  802 D--KLTA--LEAFETIPDLQEKVLQKRDEFKEKL-----QRFRENKEFNYQQFSVD 849

Qy   726 ETSVTTAQAAGDPEYLEQPSRSDFSKHLKE-ETIQIITKASHEHEDKSPETVLQSAIKLE 784
      : | : | | | | | | | | | | : | : | | | | |
Db  850 SPTVQSTE-----ETP--SSFSWYRDELEDIDI-----GDEANP-----TKND 886

Qy   785 YARLVKLAQEDTP--PETDYRLHHVVVYFIQNQ----APKKIIE-----KTLLE-- 827

```

```

      | : | | : : : | : | | : | : : |
Db    887 YLLNARLDSYSVPIAPDVETSEMCRMVSMQMNQITKIAPRKYTDSQDLLDSNLRSMVEGE 946
Qy    828 ----QFGDRNL$FDERCHNIMKVAQAKLEMIKPEEVNLEEYEEWHQDYRKFRETMYLII 883
      | : | : | | : | : : | | : | : : : :
Db    947 SGGINFINSRLGYD--IHELRSADNDVEGVYNAFIN--EYLGLVKDLHELQNSKFVVFV 1002
Qy    884 GLENFQRESYIDSLFLICAYQNNKEL-----LSKGLYRGHDEE-----LIS 925
      | : | | : : | : | | | | : | : | : |
Db    1003 GFQ-LQR-IHVPVLRVLLVRAMAVSELGIISQRANNELSGMSSNSHDKGTAMLQIALLLS 1060
Qy    926 HY 927
      | :
Db    1061 HF 1062

```

Result No.	Score	Query Match Length	DB	ID	Description
1	5410	96.9	1087	1 UBPP_HUMAN	Q9uhp3 homo sapien
2	5091	91.2	1055	1 UBPP_MOUSE	P57080 mus musculu
3	485.5	8.7	1302	1 UBPP_CAEEL	Q09931 caenorhabdi
4	323.5	5.8	1272	1 UBPP_YEAST	Q01476 saccharomyc
5	316.5	5.7	1108	1 UBPP_SCHPO	Q09879 schizosacch
6	310	5.6	1221	1 UBPP_KLULA	Q42726 kluyveromyc
7	287.5	5.2	898	1 UBPP_DROME	Q24574 drosophila
8	276	4.9	1102	1 UBPP_HUMAN	Q93009 homo sapien
9	253.5	4.5	1230	1 UBPP_YEAST	P50101 saccharomyc
10	226	4.0	499	1 UBPP_YEAST	P43593 saccharomyc
11	224	4.0	2547	1 FAFX_HUMAN	Q93008 h probable

```

RESULT 1
UBPP_HUMAN
ID UBPP_HUMAN STANDARD; PRT; 1087 AA.
AC Q9UHP3;
DT 30-MAY-2000 (Rel. 39, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE UBIQUITIN CARBOXYL-TERMINAL HYDROLASE 25 (EC 3.1.2.15) (UBIQUITIN
DE THIOLESTERASE 25) (UBIQUITIN-SPECIFIC PROCESSING PROTEASE 25)
DE (DEUBIQUITINATING ENZYME 25) (USP ON CHROMOSOME 21).
GN USP25 OR USP21.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (SHORT AND LONG ISOFORMS).
RC TISSUE=Fetal brain;
RX MEDLINE=20112753; PubMed=10644437;
RA Valero R., Marfany G., Gonzalez-Angulo O., Gonzalez-Gonzalez G.,
RA Puellas L., Gonzalez-Duarte R.;
RT "USP25, a novel gene encoding a deubiquitinating enzyme, is located in
RT the gene-poor region 21q11.2.";
RL Genomics 62:395-405(1999).
RN [2]
RP SEQUENCE FROM N.A. (SHORT ISOFORM).
RX MEDLINE=20080730; PubMed=10612803;
RA Groet J., Ives J.H., Jones T.A., Danton M., Flomen R.H., Sheer D.,
RA Hrascan R., Pavelic K., Nizetic D.;
RT "Narrowing of the region of allelic loss in 21q11-21 in squamous non-
RT small cell lung carcinoma and cloning of a novel ubiquitin-specific
RT protease gene from the deleted segment.";
RL Genes Chromosomes Cancer 27:153-161(2000).
CC -!- FUNCTION: INVOLVED IN THE UBIQUITIN-DEPENDENT PROTEOLYTIC PATHWAY
CC IN CONJUNCTION WITH THE 26S PROTEASOME (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: UBIQUITIN C-TERMINAL THIOLESTER + H(2)O =
CC UBIQUITIN + A THIOL.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A
CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.

```

CC -!- TISSUE SPECIFICITY: FOUND IN MOST ADULT AND FETAL TISSUES;  
 CC EXPRESSION IS MODERATELY HIGH IN TESTIS, PANCREAS, KIDNEY,  
 CC SKELETAL MUSCLE, LIVER, LUNG, PLACENTA, BRAIN, HEART, BUT VERY LOW  
 CC IN PERIPHERAL BLOOD, COLON, SMALL INTESTINE, OVARY, PROSTATE,  
 CC THYMUS AND SPLEEN.  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C19; ALSO KNOWN AS  
 CC FAMILY 2 OF UBIQUITIN CARBOXYL-TERMINAL HYDROLASES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF170562; AAF32263.1; -.  
 DR EMBL; AF134213; AAF24998.1; -.  
 DR MIM; 604736; -.  
 DR InterPro; IPR001394; -.  
 DR Pfam; PF00442; UCH-1; 1.  
 DR Pfam; PF00443; UCH-2; 1.  
 DR PROSITE; PS00972; UCH\_2\_1; 1.  
 DR PROSITE; PS00973; UCH\_2\_2; 1.  
 DR PROSITE; PS50235; UCH\_2\_3; 1.  
 KW Ubiquitin conjugation; Hydrolase; Thiol protease; Multigene family;  
 KW Alternative splicing.  
 FT ACT\_SITE 178 178 BY SIMILARITY.  
 FT ACT\_SITE 599 599 BY SIMILARITY.  
 FT ACT\_SITE 607 607 BY SIMILARITY.  
 FT VARSPLIC 780 811 MISSING (IN SHORT ISOFORM).  
 FT CONFLICT 544 544 K -> E (IN REF. 2).  
 SQ SEQUENCE 1087 AA; 125748 MW; 0699DDB7451B2C6C CRC64;

Query Match 96.9%; Score 5410; DB 1; Length 1087;  
 Best Local Similarity 94.9%; Pred. No. 1.5e-283;  
 Matches 1051; Conservative 1; Mismatches 3; Indels 52; Gaps 2;

```

Qy 1 MTVEQNVLQQSAAQKHQQTFLNQLREITGINDTQILQQALKDSNGNLELAVAFLTAKNAK 60
Db 1 MTVEQNVLQQSAAQKHQQTFLNQLREITGINDTQILQQALKDSNGNLELAVAFLTAKNAK 60

Qy 61 TPQQEETTYQTALPGNDRIYSVGSQADTNVIDLTGDDKDDLQRTIALSLAESNRAFRET 120
Db 61 TPQQEETTYQTALPGNDRIYSVGSQADTNVIDLTGDDKDDLQRAIALSLAESNRAFRET 120

Qy 121 GITDEEQAISRVLEASIAENKACLRTPTEVWRDSRNPYDRKRQDKAPVGLKNVGNTCWF 180
Db 121 GITDEEQAISRVLEASIAENKACLRTPTEVWRDSRNPYDRKRQDKAPVGLKNVGNTCWF 180

Qy 181 SAVIQSLFNLLEFRRLVLNYPKPSNAQDLPRNQKEHRNLFFMRELRYLFALLVGTKRKYV 240
Db 181 SAVIQSLFNLLEFRRLVLNYPKPSNAQDLPRNQKEHRNLFFMRELRYLFALLVGTKRKYV 240

Qy 241 DPSRAVEILKDAFKSNDSSQQQDVSEFTHKLLDWLEDAFQMKABEETDEEKPKNPMVELFY 300
Db 241 DPSRAVEILKDAFKSNDSSQQQDVSEFTHKLLDWLEDAFQMKABEETDEEKPKNPMVELFY 300

Qy 301 GRFLAVGVLEGKKFENTEMFGQYPLQVNGFKDLHECLEAAMIEGEIESLHSENSGKSGQE 360
Db 301 GRFLAVGVLEGKKFENTEMFGQYPLQVNGFKDLHECLEAAMIEGEIESLHSENSGKSGQE 360

Qy 361 HWFTGLPPVLTFXLSRFEFNQALGRPEKIHKNLEFPQVLYLDHYMHRNREITRIKREEIK 420
Db 361 HWFTGLPPVLTFXLSRFEFNQALGRPEKIHKNLEFPQVLYLDHYMHRNREITRIKREEIK 420

Qy 421 RLKDYLTVLQQLRLERYLSYSGGPKRFPLVDVLQYALEFASSKPVCTSPVDDIDASSPPSG 480
Db 421 RLKDYLTVLQQLRLERYLSYSGGPKRFPLVDVLQYALEFASSKPVCTSPVDDIDASSPPSG 480

```



Qy	481	SIP	SQ	TLP	STTE	QQ	GAL	SS	ELP	ST	SP	SS	VAA	IS	SR	SV	IHK	PFT	Q	S	R	I	P	P	D	L	P	M	H	P	A	P	R	H	I	540																										
Db	481	SIP	SQ	TLP	STTE	QQ	GAL	SS	ELP	ST	SP	SS	VAA	IS	SR	SV	IHK	PFT	Q	S	R	I	P	P	D	L	P	M	H	P	A	P	R	H	I	540																										
Qy	541	TEE	EL	SV	LE	SC	LR	WR	TE	I	EN	DR	DL	Q	ES	I	S	R	I	H	R	T	I	E	L	M	Y	S	D	K	S	M	I	Q	V	P	Y	R	L	H	A	V	L	V	H	E	600															
Db	541	TEE	EL	SV	LE	SC	LR	WR	TE	I	EN	DR	DL	Q	ES	I	S	R	I	H	R	T	I	E	L	M	Y	S	D	K	S	M	I	Q	V	P	Y	R	L	H	A	V	L	V	H	E	600															
Qy	601	Q	A	N	A	G	H	Y	W	A	I	F	D	H	R	E	S	R	W	M	K	Y	N	D	I	A	V	T	K	S	S	W	E	E	L	V	R	D	S	F	G	G	Y	R	N	A	S	A	C	L	M	Y	I	N	D	K	A	Q	660			
Db	601	Q	A	N	A	G	H	Y	W	A	I	F	D	H	R	E	S	R	W	M	K	Y	N	D	I	A	V	T	K	S	S	W	E	E	L	V	R	D	S	F	G	G	Y	R	N	A	S	A	C	L	M	Y	I	N	D	K	A	Q	660			
Qy	661	F	L	I	Q	E	E	F	N	K	E	T	G	Q	P	L	V	G	I	E	T	L	P	P	D	L	R	D	F	V	E	E	D	N	Q	R	F	E	K	E	L	E	W	D	A	Q	L	A	Q	K	A	L	Q	E	K	L	L	A	S	Q	720	
Db	661	F	L	I	Q	E	E	F	N	K	E	T	G	Q	P	L	V	G	I	E	T	L	P	P	D	L	R	D	F	V	E	E	D	N	Q	R	F	E	K	E	L	E	W	D	A	Q	L	A	Q	K	A	L	Q	E	K	L	L	A	S	Q	720	
Qy	721	K	L	R	E	S	E	T	S	V	T	T	A	Q	A	G	D	P	E	Y	L	E	Q	P	S	R	S	D	F	S	K	H	L	K	E	E	T	I	Q	I	T	K	A	S	H	E	H	E	D	K	S	P	E	T	V	L	Q	S	-	779		
Db	721	K	L	R	E	S	E	T	S	V	T	T	A	Q	A	G	D	P	E	Y	L	E	Q	P	S	R	S	D	F	S	K	H	L	K	E	E	T	I	Q	I	T	K	A	S	H	E	H	E	D	K	S	P	E	T	V	L	Q	S	I	780		
Qy	780	-----AIKLEYARLVKLAQEDTPPETDYRLHHVV																									808																																			
Db	781	-----AIKLEYARLVKLAQEDTPPETDYRLHHVV																									840																																			
Qy	809	V	F	I	Q	N	A	P	K	K	I	E	K	T	L	L	E	Q	F	G	D	R	N	L	S	F	D	E	R	C	H	N	I	M	K	V	A	Q	A	K	L	E	M	I	K	P	E	E	V	N	L	E	E	Y	E	W	H	868				
Db	841	V	F	I	Q	N	A	P	K	K	I	E	K	T	L	L	E	Q	F	G	D	R	N	L	S	F	D	E	R	C	H	N	I	M	K	V	A	Q	A	K	L	E	M	I	K	P	E	E	V	N	L	E	E	Y	E	W	H	900				
Qy	869	Q	D	Y	R	K	F	R	E	T	T	M	Y	L	I	I	G	L	E	N	F	Q	R	E	S	Y	I	D	S	L	L	F	L	I	C	A	Y	Q	N	N	K	E	L	L	S	K	G	L	Y	R	G	H	D	E	E	L	I	S	H	Y	928	
Db	901	Q	D	Y	R	K	F	R	E	T	T	M	Y	L	I	I	G	L	E	N	F	Q	R	E	S	Y	I	D	S	L	L	F	L	I	C	A	Y	Q	N	N	K	E	L	L	S	K	G	L	Y	R	G	H	D	E	E	L	I	S	H	Y	960	
Qy	929	R	E	C	L	L	I	L	N	L	K	R	K	Q	P	I	L	F	F	L	H	C	I	K	L	N	E	Q	A	A	E	L	F	E	S	G	E	D	R	E	V	N	N	G	L	I	M	N	E	F	I	V	P	F	L	P	L	988				
Db	961	R	E	C	L	L	-----	-----KLNEQAALFESGEDREVNNGLIIMNEFIVPFLPL																				1000																																		
Qy	989	L	L	V	D	E	M	E	E	K	D	I	L	A	V	E	D	M	R	N	R	W	C	S	Y	L	G	Q	E	M	E	P	H	L	Q	E	K	L	T	D	F	L	P	K	L	L	D	C	S	M	E	I	K	S	F	H	E	P	P	K	L	1048
Db	1001	L	L	V	D	E	M	E	E	K	D	I	L	A	V	E	D	M	R	N	R	W	C	S	Y	L	G	Q	E	M	E	P	H	L	Q	E	K	L	T	D	F	L	P	K	L	L	D	C	S	M	E	I	K	S	F	H	E	P	P	K	L	1060
Qy	1049	P	S	Y	S	T	H	E	L	C	E	R	F	A	R	I	M	L	S	L	S	R	T	P	A	D	G	R	1075																																	
Db	1061	P	S	Y	S	T	H	E	L	C	E	R	F	A	R	I	M	L	S	L	S	R	T	P	A	D	G	R	1087																																	

RESULT 2

UBPP\_MOUSE

ID UBPP\_MOUSE STANDARD; PRT; 1055 AA.

AC P57080;

DT 01-OCT-2000 (Rel. 40, Created)

DT 01-OCT-2000 (Rel. 40, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE UBIQUITIN CARBOXYL-TERMINAL HYDROLASE 25 (EC 3.1.2.15) (UBIQUITIN THIOLESTERASE 25) (UBIQUITIN-SPECIFIC PROCESSING PROTEASE 25)

DE (DEUBIQUITINATING ENZYME 25) (MUSP25).

GN USP25.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Kidney;

RX MEDLINE=20112753; PubMed=10644437;

RA Valero R., Marfany G., Gonzalez-Angulo O., Gonzalez-Gonzalez G.,

RA Puellas L., Gonzalez-Duarte R.;

RT "USP25, a novel gene encoding a deubiquitinating enzyme, is located in the gene-poor region 21q11.2.";

RL Genomics 62:395-405(1999).

CC -!- FUNCTION: INVOLVED IN THE UBIQUITIN-DEPENDENT PROTEOLYTIC PATHWAY

CC IN CONJUNCTION WITH THE 26S PROTEASOME (BY SIMILARITY).

```

CC  -!- CATALYTIC ACTIVITY: UBIQUITIN C-TERMINAL THIOLESTER + H(2)O =
CC      UBIQUITIN + A THIOL.
CC  -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN TESTIS. ALSO FOUND IN
CC      BRAIN, SKELETAL MUSCLE, LIVER AND HEART.
CC  -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C19; ALSO KNOWN AS
CC      FAMILY 2 OF UBIQUITIN CARBOXYL-TERMINAL HYDROLASES.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AF170563; AAF32264.1; -.
DR  MGD; MGI:1353655; Usp25.
DR  PROSITE; PS00972; UCH_2_1; 1.
DR  PROSITE; PS00973; UCH_2_2; 1.
DR  PROSITE; PS50235; UCH_2_3; 1.
KW  Ubiquitin conjugation; Hydrolase; Thiol protease; Multigene family.
FT  ACT_SITE    178    178      BY SIMILARITY.
FT  ACT_SITE    600    600      BY SIMILARITY.
FT  ACT_SITE    608    608      BY SIMILARITY.
SQ  SEQUENCE    1055 AA; 121426 MW; CDFDFD1052745D70 CRC64;

```

```

Query Match          91.2%; Score 5091; DB 1; Length 1055;
Best Local Similarity 91.5%; Pred. No. 2.1e-266;
Matches 985; Conservative 29; Mismatches 40; Indels 22; Gaps 3;

```

```

Qy   1 MTVEQNVLQSSAAQKHQQTFLNQLREITGINDTQILQQALKDSNGNLELAVAFLTAKNAK 60
      |||
Db   1 MTVEQNVLQSSAAQKHQQTFLNQLREITGINDAQILQQALKDSNGNLELAVAFLTAKNAK 60

Qy  61 TPQQEETTYQTALPGNDRIYSVGSQADTNVIDLTGDDKDDLQRTIALSLAESNRAFRET 120
      || |||
Db  61 TPPQEETGYQTALPGNDRIYSVGSQADANVIDLTGDDKDDLQRAIALSLAESNRAFRET 120

Qy  121 GITDEEQAISRVLEASIAENKACLKRTPTFVWRDSRNPYDRKRQDKAPVGLKNVGNCTCF 180
      |||
Db  121 GITDEEQAISRVLEASIAENKACLKRTPIFVWRDSRNPYDRKRQEKAPVGLKNVGNCTCF 180

Qy  181 SAVIQSLFNLLEFRRLLVNLKPPSNAQDLPRNQKEHRNLPFMRELRYLFALLVGTKRKYV 240
      |||
Db  181 SAVIQSLFNLLEFRRLLVNLKPPSNAQDLPRNQKEHRNLPFMRELRYLFALLVGTKRKYV 240

Qy  241 DPSRAVEILKDAFKNDSQQQDVSEFTHKLLDWLEDAFQMKAEETDEEKPKNPMVELFY 300
      |||
Db  241 DPSRAVEILKDAFKNDSQQQDVSEFTHKLLDWLEDAFQMKAEETDEEKPKNPMVELFY 300

Qy  301 GRFLAVGVLEGGKFENTEMFGQYPLQVNGFKDLHECLEAAMIEGEIESLHSDNSGKSGQE 360
      |||:|
Db  301 GRFLAMGVLEGGKFENTEMFGQYPLQVNGFKDLHECLEAAMIEGEIESLHSDNSGKSGQE 360

Qy  361 HWFTGLPPVLTFLXLSRFEFNQALGRPEKIHNKLEFPQVLYLDRYMHRNREITRIKREEIK 420
      |||
Db  361 HWFTGLPPVLTFLXLSRFEFNQALGRPEKIHNKLEFPQVLYLDRYMHRNREITRIKREEIK 420

Qy  421 RLKDYLTVLQQRLELYLSYSGGPKRFPLVDVLQYALEFASSKPVCTSPVDDIDASSPPSG 480
      |||
Db  421 RLKDYLTVLQQRLELYLSYSGGPKRFPLVDVLQYALEFASSKPVCTSPVDDIDASSSASG 480

Qy  481 SIPSQTLPTSTTEQQGALSSELP-STSPSSVAAISSRSVIHKPFTQSRIPPDLPMPHAPRH 539
      :||:|
Db  481 PLPSQSLPTSTTEQQGLCASDLPSSSPASGAALPLRSVIHKPFTQSRIPPDLPMPHAPRH 540

Qy  540 ITEEELSVLESCLHRWRTEIENDTRDLQESISRIHRTIELMYSDKSMIQVPYRLHAVLVH 599
      |||
Db  541 ITEEELCVLESCLHRWRTEIENDTRDLQESISRIHRTIELMYSDKSMIQVPYRLHAVLVH 600

```

Qy 600 EGQANAGHYWAYIFDHRESRWMKYNDIAVTKSSWEELVRDSFGGYRNASAYCLMYINDKA 659  
 Db 601 EGQANAGHYWAYIFDHRESRWMKYNDIAVTKSSWEELVRDSFGGYRNASAYCLMYIDDKA 660

Qy 660 QFLIQEEFNKETGQPLVGIETLPPDLRDFVEEDNQRFKEKELEEWDAQLAQKALQEKLAS 719  
 Db 661 QFLIQEEFNKETGQALVGMETLLPDLRDFVEEDNQRFKEKELEEWDTQLAQRSLQEKLAA 720

Qy 720 QKLRSETS VTTAQAAGDPEYLEQPSRSDFSKHLKEETIQIITKASHEHEDKSPETVLQS 779  
 Db 721 PKLREAEASATTAQAGG-ADYLEQPSRSDLSKHWKEETLRVIKASHDLEDKGPETVLQS 779

Qy 780 AIKLEYARLVKLAQEDTPPETDYRLHHVVVFYFQIQAPKKIIEKTLLEQFGDRNLSFDER 839  
 Db 780 AIKLEYSRLVKLAQEDTPPETDYRLHHVLFYFQIQAPKKIIEKTLLEQFGDRNLSFDER 839

Qy 840 CHNIMKVAQAKLEMIKPEEVNLEEYEEWHQDYRKFRETTMYLIIGLENFQRESYIDSLLF 899  
 Db 840 CLNIMKVAQAKLEMIKPEEVNLEEYEEWHADYKKFRETTMYLITGLENFQRESYIDSLLF 899

Qy 900 LICAYQNNKELLSKGLYRGHDEELISHYRRECLLNLKRKQKPILFFFLHCICKLNEQA 959  
 Db 900 LLCAYQNNKELLSKGPYRGHDGELISHYRRECLL-----KLNEQA 939

Qy 960 AELFESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEEKDILAVEDMRNRWCSYLGQEMEP 1019  
 Db 940 AELFESGEDGDVNNGLIIMNEFIVPFLPLLLVDDMEEKDILAVEDMRNRWCSYLGQEMEA 999

Qy 1020 HLQEKLTDFLPKLLDCSMEIKSFHEPPKLPSSYTHELCERFARIMLSLSRTPADGR 1075  
 Db 1000 NLQEKLTDFLPKLLDCSTEIKGFHEPPKLPSSYSAHELCEFRARIMLSLSRTPADGR 1055

Result No.	Score	Query Match	Length	DB	ID	Description
1	3816	68.4	737	4	Q9H9W1	Q9h9w1 homo sapien
2	2721	48.8	523	4	Q9HA22	Q9ha22 homo sapien
3	1875	33.6	757	4	Q9P213	Q9p213 homo sapien
4	303	5.4	1124	10	Q9SF08	Q9sf08 arabidopsis
5	297.5	5.3	1116	10	Q9FPT1	Q9fpt1 arabidopsis
6	297	5.3	1126	10	Q9FG10	Q9fg10 arabidopsis
7	296	5.3	1248	5	Q9VRP1	Q9vrp1 drosophila
8	288.5	5.2	599	10	Q9FU99	Q9fu99 arabidopsis
9	287	5.1	1130	5	O45624	O45624 caenorhabdi
10	287	5.1	1133	5	O45623	O45623 caenorhabdi
11	283.5	5.1	1141	3	Q9P3U0	Q9p3u0 schizosacch

# RESULT 1

## Q9H9W1

ID Q9H9W1 PRELIMINARY; PRT; 737 AA.  
 AC Q9H9W1;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
 DE CDNA FLJ12512 FIS, CLONE NT2RM2001730, WEAKLY SIMILAR TO PROBABLE  
 DE UBIQUITIN CARBOXYL-TERMINAL HYDROLASE K02C4.3 (EC 3.1.2.15).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,  
 RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,  
 RA Nakamura Y., Nagahari K., Masuho Y., Sasaki N.;  
 RT "NEDO human cDNA sequencing project."  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AK022574; BAB14107.1; -.  
SQ SEQUENCE 737 AA; 85842 MW; D660BB2A18077E10 CRC64;

Query Match 68.4%; Score 3816; DB 4; Length 737;  
Best Local Similarity 97.1%; Pred. No. 6.4e-221;  
Matches 735; Conservative 0; Mismatches 2; Indels 20; Gaps 1;

```
Qy 319 MFGQYPLQVNGFKDLHECLEAAMIEGEIESLHSENSGKSGQEHWFTGLPPVLTFFXLSRFE 378
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 MFGQYPLQVNGFKDLHECLEAAMIEGEIESLHSENSGKSGQEHWFTGLPPVLTFFELSRFE 60

Qy 379 FNQALGRPEKIHNKLEFPQVLYLDRYMHRNREITRIKREEIKRLKDYLTVLQQRLERYLS 438
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 FNQALGRPEKIHNKLEFPQVLYLDRYMHRNREITRIKREEIKRLKDYLTVLQQRLERYLS 120

Qy 439 YGSGPKRFPPLVDVLQYALEFASSKPVCTSPVDDIDASSPPSGSIPSQTLPTSTTEQQGALS 498
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 YGSGPKRFPPLVDVLQYALEFASSKPVCTSPVDDIDASSPPSGSIPSQTLPTSTTEQQGALS 180

Qy 499 SELPSTSPSSVAAISSRSVIHKPFTQSRIPDLPMHPAPRHITTEELSULESCLHRWRTE 558
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 SELPSTSPSSVAAISSRSVIHKPFTQSRIPDLPMHPAPRHITTEELSULESCLHRWRTE 240

Qy 559 IENDTRDLQESISRIHRTIELMYSKSMIQVPYRLHAVLVHEGQANAGHYWAYIFDHRES 618
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 IENDTRDLQESISRIHRTIELMYSKSMIQVPYRLHAVLVHEGQANAGHYWAYIFDHRES 300

Qy 619 RWMKYNDIAVTKSSWEELVRDSFGGYRNASAYCLMYINDKAQFLIQEEFNKETGQPLVGI 678
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 RWMKYNDIAVTKSSWEELVRDSFGGYRNASAYCLMYINDKAQFLIQEEFNKETGQPLVGI 360

Qy 679 ETLPDLRDFVEEDNQRFKELEEWDAQLAQKALQEKLLASQKLRESETSVTTAQAGDP 738
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 ETLPDLRDFVEEDNQRFKELEEWDAQLAQKALQEKLLASQKLRESETSVTTAQAGDP 420

Qy 739 EYLEQPSRSDFSKHLKEETIQIITKASHEHEDKSPETVLQSAIKLEYARLVKLAQEDTPP 798
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 421 EYLEQPSRSDFSKHLKEETIQIITKASHEHEDKSPETVLQSAIKLEYARLVKLAQEDTPP 480

Qy 799 ETDYRLHHVVVYFIQNQAPKKIIEKTLLEQFGDRNLSFDERCHNIMKVAQAKLEMIKPEE 858
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 481 ETDYRLHHVVVYFIQNQAPKKIIEKTLLEQFGDRNLSFDERCHNIMKVAQAKLEMIKPEE 540

Qy 859 VNLEEEYEWHDYRKFRETTMVLIIGLENFQRESYIDSLFLICAYQNNKELLSKGLYRG 918
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 541 VNLEEEYEWHDYRKFRETTMVLIIGLENFQRESYIDSLFLICAYQNNKELLSKGLYRG 600

Qy 919 HDEELISHYRRECLLILNLKRKQKPIILFFFLHCKLNEQAAELFESGEDREVNNGLIIM 978
      |||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db 601 HDEELISHYRRECLL-----KLNEQAAELFESGEDREVNNGLIIM 640

Qy 979 NEFIVPFLPLLLVDEMEEKDILAVEDMRNRWCSYLGQEMEPHLQEKLTDFLPKLLDCSME 1038
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 641 NEFIVPFLPLLLVDEMEEKDILAVEDMRNRWCSYLGQEMEPHLQEKLTDFLPKLLDCSME 700

Qy 1039 IKSFPHEPPKLPSYSTHELCEFRARIMLSLSRTPADGR 1075
      ||||||||||||||||||||||||||||||||||||||||||||
Db 701 IKSFPHEPPKLPSYSTHELCEFRARIMLSLSRTPADGR 737
```